

(19) BUNDESREPUBLIK

DEUTSCHLAND



DEUTSCHES

PATENTAMT

(12) Offenlegungsschrift

(10) DE 197 08 265 A 1

(51) Int. Cl. 6:

G 06 F 17/30

G 06 K 9/46

DE 197 08 265 A 1

(21) Aktenzeichen: 197 08 265.3
(22) Anmeldetag: 28. 2. 97
(23) Offenlegungstag: 4. 9. 97

(30) Unionspriorität:

609641 01.03.96 US

(71) Anmelder:

Ricoh Co., Ltd., Tokio/Tokyo, JP

(74) Vertreter:

Schwabe, Sandmair, Marx, 81677 München

(72) Erfinder:

Cullen, John, Menlo Park, Calif., US; Hull, Jonathan, Menlo Park, Calif., US; Hart, Peter, Menlo Park, Calif., US

Prüfungsantrag gem. § 44 PatG ist gestellt

(54) Blättern in einer Bilddatenbank und Abfrage der Bilddaten, indem eine Strukturanalyse verwendet wird

(57) Verfahren und Apparat zum Abfragen einer Dokument-Bilddatenbank, basierend auf einer Struktur, d. h. auf analytisch unterscheidbaren Mustern in den Dokument-Bildern der Datenbank. Gemäß der Erfindung kann eine Dokumentbilddatenbank für Dokumente mit einer bestimmten Struktur auf eine Vielfalt von Weisen durchblättert werden. Zum Beispiel kann ein Benutzer ein Beispiel-Dokumentbild mit einem Erscheinungsbild, das dem gewünschten Dokument ähnelt, eingeben. Alternativ kann der Benutzer eine einfache Schnittstelle verwenden, um ein synthetisches Dokument bzw. künstliches Dokument, basierend auf einer Auswahl weniger Kategorien festzulegen. Das synthetische Dokument würde dann als ein Beispiel für eine Suche dienen. Oder der Benutzer kann eine graphische Schnittstelle verwenden, um genauer ein Beispiel für eine Suche festzulegen. Somit liefert die Kenntnis des Benutzers von dem allgemeinen Erscheinungsbild des gewünschten Dokuments oder der Dokumente die Basis für die Suche.

DE 197 08 265 A 1

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

BUNDESDRUCKEREI 07.97 702 036/516

22/23

Beschreibung

Die vorliegende Erfindung betrifft ein Verfahren und einen Apparat zum Durchstöbern einer Bilddatenbank bzw. zum Blättern in einer Bilddatenbank und genauer ein Verfahren und einen Apparat zum Durchstöbern einer Bilddatenbank bzw. zum Blättern in einer Bilddatenbank, und zwar basierend auf einer bekannten Struktur bzw. bekannten Textur eines gewünschten Dokuments.

Mit der weiten Verbreitung preisgünstiger Permanent-Elektronikspeichervorrichtungen und von Scannern zum elektronischen Erfassen von Dokumentbildern bzw. Vorlagenbildern, wird die Erzeugung und Speicherung von großen Dokumentbild-Datenbanken möglich. Mögliche Verwendungen für derartige Bilddatenbanken stellen elektronische Bibliotheken, medizinische Dokumente, Verwaltungsaufzeichnungen, Nachrichtenclips, Pläne, Verwaltungsformulare, Handbücher, usw. dar. Im allgemeinen ist es nun weniger teuer, ein Dokument abzutasten und es auf einem elektronischen Speichermedium zu speichern, als das Dokument physisch zu speichern.

Bei derartigen großen Dokument-Bilddatenbanken ergibt sich jedoch das Problem, gewünschte Dokumente aus der Datenbank zu identifizieren und wiederzugewinnen. Eine bekannte Technik zum Abfragen einer Vorlagendatenbank besteht darin, nach einem Textstring bzw. nach einer Textzeichenkette oder einer Kombination von Textstrings bzw. Textzeichenketten zu suchen, die wahrscheinlich in einem gewünschten Dokument gefunden werden. Was das abgetastete Dokument angeht, beruht diese Technik auf einer genauen optischen Zeichenerkennung (OCR bzw. "optical character recognition"), die nicht immer machbar sein kann. Ebenso ist es häufig schwierig, einen Textstring oder Strings zu erzeugen, die sowohl in einem gewünschten Dokument gefunden werden als auch im ausreichenden Maße das gewünschte Dokument von anderen unterscheiden.

Häufig weiß der Benutzer etwas darüber, wie ein gewünschtes Dokument aussieht. Es wäre von Vorteil, diese Information beim Abfragen einer Dokument-Bilddatenbank zu nutzen.

Die vorliegende Erfindung stellt ein Verfahren und einen Apparat zum Abfragen einer Dokument-Bilddatenbank dar, und zwar basierend auf einer Struktur bzw. einem inneren Aufbau und analytisch unterscheidbare Muster in den Dokument-Bildern der Datenbank. Gemäß der Erfindung kann eine Dokument-Bilddatenbank nach Dokumenten mit einer bestimmten Struktur bzw. einem bestimmten inneren Aufbau auf eine Vielfalt von Weisen durchstöbert bzw. durchsucht werden. Zum Beispiel kann ein Benutzer ein Beispiel-Dokumentbild eingeben, das ein zu dem gewünschten Dokument ähnliches Erscheinungsbild aufweist. Alternativ kann der Benutzer eine einfache Schnittstelle verwenden, um ein synthetisches bzw. künstliches Dokument zu definieren, und zwar basierend auf der Auswahl weniger Kategorien. Das synthetische Dokument würde dann als ein Beispiel für eine Suche dienen. Oder der Bediener kann eine graphische Schnittstelle verwenden, um genauer ein Beispiel für eine Suche festzulegen. Somit stellt die Kenntnis des Benutzers über das allgemeine Erscheinungsbild des gewünschten Dokuments oder der Dokumente (und insbesondere deren Umwandlung in physikalische Werte bzw. elektrische Signale, die von der Datenbank bzw. einem Computer verarbeitbar sind,) die

Grundlage für die Suche dar.

Wenn einmal der Benutzer die Suche gestartet hat, wird ein Dokumentbild oder werden mehrere Dokumentbilder ausgewählt, und zwar basierend auf der Ähnlichkeit ihrer Struktur oder ihres inneren Aufbaus mit dem Beispiel. Diese Bilder können in einer Piktogrammgestalt bzw. in einer Ikongestalt auf einem Papier dargestellt oder gedruckt werden. Der Benutzer kann ein weiteres Durchblättern bzw. Durchstöbern auslösen, indem ein dargestelltes Dokument als ein Suchschlüssel für eine neue Suche ausgewählt wird. Bei einer Ausführungsform werden Dokumentbilder in der Datenbank zusammen gruppiert bzw. geclustert und wenn eine Abfrage durchgeführt wird, werden Darstellungen für die Gruppen bzw. Cluster dargestellt, die Bilder beinhalten, die bezüglich ihrer Struktur dem Suchschlüssel ähneln.

Bei einer Ausführungsform werden Seiten mit einem ähnlichen Erscheinungsbild identifiziert, wenn der Benutzer die Anzeige mehrerer Seiten eines Dokuments verlangt. Bezuglich der Seiten, die dahingehend identifiziert wurden, daß sie ein ähnliches Erscheinungsbild aufweisen, werden Bilder der Seiten einander überlagert und leicht versetzt dargestellt.

In Übereinstimmung mit einem Aspekt der Erfindung liegt die Grundlage für Dokument-Bildsuchvorgänge in einer Analyse der Struktur bzw. des inneren Aufbaus eines Bildes. Ein Merkmalsvektor wird für jedes Bild extrahiert. Eine Ähnlichkeit zwischen Bildern wird auf einer Abstandsmeßikte gegründet, wie sie für die Merkmalsvektoren paßt bzw. gilt. Bei der bevorzugten Ausführungsform weist der Merkmalsvektor 80 Elemente auf und gründet auf vier unterschiedlichen Arten und Weisen der Bildverarbeitung. Die ersten 20 Elemente basieren auf einem Histogramm von verbundenen Komponentengrößen bzw. Verbindungskomponentengrößen ("connected component sizes"), und zwar durchgehend für das Bild. Verbundene Komponenten ("connected components") werden im folgenden auch Verbindungsdimensionen genannt. Die zweiten 20 Elemente werden aus einer Übersicht bzw. Vermessung von Interessenpunkten bzw. Punkten von Interesse über das ganze Dokument bzw. durchgehend bezüglich des Dokuments erhalten. Die dritten 20 Elemente werden von einem vertikalen Projektionshistogramm verbundener Bestandteile bzw. Komponenten erhalten. Der letzte Satz von 20 Elementen wird erhalten, indem das Dokument in 20 Zellen aufgeteilt wird und die verbundenen Bestandteile bzw. Komponenten in jeder Zelle aufsummiert werden.

Bei einer Ausführungsform werden Suchverfahren, basierend auf einer Kombination von Bildstruktur und Text verwendet, um eine Bilddatenbank abzufragen, und zwar dort, wo eine optische Zeichenerkennung (OCR) mit den Bildern der Datenbank durchgeführt wurde. Durch die Kombination der beiden Datentypen kann Text verwendet werden, um nach den OCR-Daten, die dem Bild zugeordnet sind, zu suchen, und die Strukturdaten können für eine Suche, basierend auf dem allgemeinen Erscheinungsbild des Bildes verwendet werden.

Die vorliegende Anmeldung basiert auf der prioritätsbegründenden US-Patentanmeldung S.N.08/609,641, angemeldet am 1. März 1996, deren Offenbarung hiermit mit aufgenommen wird.

Ein weiteres Verständnis der Natur und der Vorteile der Erfindungen hierin kann realisiert werden, indem auf die verbleibenden Abschnitte der Beschreibung und

der beigefügten Zeichnungen bezogenommen wird. Dabei können Merkmale verschiedener Ausführungsformen untereinander kombiniert werden.

Fig. 1 zeigt ein Computersystem, das zur Realisierung der vorliegenden Erfindung geeignet ist.

Fig. 2 zeigt ein Flußdiagramm auf obersten Niveau, das dem Betrieb des Dokument-Bilddatenbank-Abfragesystems der vorliegenden Erfindung beschreibt.

Fig. 3 zeigt einen Benutzer-Schnittstellen-Bildschirm zum Entwickeln eines synthetischen Dokumentenbildes, basierend auf Kategorieselektionen als ein Beispiel, um eine Suche in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung zu starten.

Fig. 4 zeigt einen Benutzer-Schnittstellen-Bildschirm, bei dem ein Benutzer graphische Werkzeuge verwenden kann, um ein Beispiel-Dokumentenbild zu entwickeln, um eine Suche in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung zu starten.

Fig. 5 zeigt, wie Suchergebnisse für eine Bilddatenbank mit Clustern bzw. Gruppen in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung gezeigt werden können.

Fig. 6 zeigt, wie Bilder mehrerer Seiten eines Berichts in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung gezeigt werden können.

Fig. 7A zeigt die Elemente eines Merkmalvektors, der aus einem Dokumentenbild in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung extrahiert wird.

Fig. 7B zeigt ein verbundenes Komponenten-Größen-Histogramm, wie es innerhalb eines Merkmalvektors in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung eingebracht werden würde.

Fig. 7C zeigt, wie Interessen-Operator-Information, wie sie innerhalb eines Merkmalvektors in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung eingebaut bzw. eingebunden werden würde, arbeitet, um Unterscheidungen zwischen Dokumenten mit unterschiedlichen Font-Größen durchzuführen.

Fig. 7D zeigt ein Vertikalprojektions-Verbindungs-komponentenhistogramm, wie es innerhalb eines Merkmalvektors in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung eingebracht bzw. eingebunden werden würde.

Fig. 7E zeigt, wie die Verteilung von verbundenen Komponenten bzw. Verbindungskomponenten unter Gitterzellen, wie sie innerhalb eines Merkmalvektors in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung inkorporiert werden würden, arbeitet, um unter Dokumenten mit unterschiedlichen Anordnungen von Komponenten zu unterscheiden.

Fig. 8 zeigt ein repräsentatives Beispiel-Dokumentenbild, das zur Suche in Übereinstimmung mit der vorliegenden Erfindung verwendet wird, und Suchergebnisse.

System, das zur Implementation der vorliegenden Erfindung geeignet ist

Fig. 1 zeigt ein Basis-Untersystem eines Computersystems, das zur Verwendung mit der vorliegenden Erfindung geeignet ist. In Fig. 1 beinhaltet das Computersystem 10 einen Bus 12, der Haupt-Untersysteme, wie zum Beispiel eine zentrale Verarbeitungseinheit bzw. einen Hauptprozessor 14, einen Systemspeicher 16, eine Eingabe/Ausgabe- (I/O bzw. "input/output") Steuereinrichtung 18, eine externe Vorrichtung, wie zum Beispiel einen Drucker 20 über einen Parallelport 22, einen Anzei-

gebildschirm 24 über einen Anzeigeadapter 26, einen seriellen Port 28, eine Tastatur 30, ein Festplattenlaufwerk 32 und ein Diskettenlaufwerk 33, das arbeitet, um eine Diskette 33A aufzunehmen, verbindet. Viele andere Vorrichtungen können angeschlossen werden, wie zum Beispiel eine Abtastvorrichtung 34, das über eine externe Schnittstelle 36 angeschlossen ist, eine Maus 38, die über ein serielles Port 28 angeschlossen ist, und einen Berührungsbildschirm bzw. Touchscreen 40, der direkt angeschlossen ist. Viele andere Vorrichtungen und Untersysteme (nicht gezeigt) können auf eine ähnliche Art und Weise angeschlossen werden. Ebenso ist es nicht für alle in der Fig. 1 gezeigten Vorrichtungen erforderlich, daß sie vorhanden sind, um die vorliegende Erfindung in die Praxis umzusetzen, wie im folgenden diskutiert wird. Die Vorrichtungen und Untersysteme können miteinander auf eine Art und Weise verbunden werden, die von jener verschieden ist, die in Fig. 1 gezeigt ist. Der Betrieb eines Computersystems, wie zum Beispiel jener, der in Fig. 1 gezeigt ist, ist in der Fachwelt gut bekannt und wird in der vorliegenden Anmeldung nicht detailliert diskutiert. Ein Quellcode, um die vorliegende Erfindung zu realisieren, kann betriebsmäßig in einem Systemspeicher 16 abgelegt oder auf einem Speichermedium, wie zum Beispiel einer Festplatte oder einer Diskette 33A gespeichert werden. Eine Bilddatenbank kann ebenso auf einer Festplatte 32 gespeichert werden.

Ein Anzeigebildschirm 24 ähnelt jenem, der bei Standardcomputern, wie zum Beispiel Personalcomputern oder Workstations verwendet wird, die einen Kathodenstrahlröhren(CRT)-Bildschirm oder einen Monitor verwenden. Verschiedene Formen von Bediener-Eingabevorrichtungen können zusammen mit der vorliegenden Erfindung verwendet werden. Zum Beispiel eine Maus-Eingabevorrichtung, die es einem Benutzer erlaubt, einen Zeiger, der auf einem Anzeigebildschirm dargestellt wird, in Übereinstimmung mit den Handbewegungen eines Benutzers zu bewegen, stellt eine Standard-Benutzereingabevorrichtung dar. Eine Maus beinhaltet üblicherweise eine oder mehrere Tasten auf seiner Oberfläche, so daß der Benutzer auf ein Objekt auf dem Bildschirm zeigen kann, indem die Maus bewegt wird und das Objekt ausgewählt wird, oder indem auf andere Weise das Objekt aktiviert wird, indem eine oder mehrere Tasten auf der Maus gedrückt werden. Alternativ erlaubt ein Berührungsbildschirm bzw. ein Touchscreen einen Benutzer, auf Objekte auf dem Bildschirm zu zeigen, um ein Objekt auszuwählen, und um das ausgewählte Objekt zu bewegen, indem auf eine zweite Position auf dem Bildschirm gezeigt wird. Verschiedene Tasten und Steuerungen können auf dem Bildschirm gezeigt werden, um sie zu aktivieren, indem die Maus oder der Touchscreen verwendet wird. Bei einem Festplattenlaufwerk 32 kann es sich um ein Festplattenlaufwerk oder um ein optisches Laufwerk oder um irgendein Medium handeln, das zum Speichern einer Datenbank von Dokumentenbildern geeignet ist.

Überblick über das Suchverfahren

Die vorliegende Erfindung stellt ein Dokument-Bilddatenbank-Suchsystem bereit, das auf der Verwendung von Beispieldokumenten bzw. als Beispiel dienenden Dokumenten als Schlüssel für die Suche basiert. Durch Wechselwirkung mit dem System der vorliegenden Erfindung auf eine einer Anzahl von Weisen entwickelt der Benutzer ein Beispielbild. Vorzugsweise findet das System, basierend auf einer Struktur des Beispielbildes

eine oder mehrere Übereinstimmungen und zeigt sie als Suchergebnisse an. Der Benutzer kann dann ein übereinstimmendes Bild als einen Schlüssel für eine weitere Suche in der Datenbank auswählen.

Fig. 2 zeigt ein Flußdiagramm auf oberstem Niveau, das dem Betrieb des Dokument-Bilddatenbank-Abfragesystems der vorliegenden Erfindung beschreibt. Im Schritt 202 wird ein Beispielbild entwickelt. Zum Beispiel kann der Benutzer ein bekanntes Beispielbild aus einem Diskettenlaufwerk 32 wiederfinden. Oder der Benutzer kann ein Bild einscannen, indem eine Abtastvorrichtung 34 verwendet wird. Andere Techniken zum Entwickeln eines Beispielbildes werden unter Bezugnahme auf Fig. 3 und 4 beschrieben.

Das Beispielbild, das im Schritt 202 entwickelt wird, dient als Grundlage für die Suche. Die Suche beruht auf einem Vergleich von Bildmerkmalen, wie sie über Bereiche des Dokuments gewonnen bzw. berechnet werden. Die Wirkung liegt darin, daß Dokumentbilder gefunden werden, die ein zu dem Dokumentbild ähnliches allgemeines Erscheinungsbild aufweisen. Dementsprechend wird dem Schritt 204 eine Bildmerkmalsinformation über das Beispiel-Dokumentbild erhalten. Vorzugsweise nimmt diese Bildmerkmalsinformation die Gestalt eines Bildmerkmalvektors ein, dessen Zusammensetzung unter Bezugnahme auf die Fig. 7A-7E diskutiert wird.

Die Datenbasis beinhaltet vorzugsweise auf ähnliche Weise erhaltene Bildmerkmalsinformationen für die Datenbankbilder. Im Schritt 206 wird die Bildmerkmalsinformation, die aus dem Beispiel-Dokumentenbild erhalten wird, als ein Schlüssel für eine Suche in der Datenbank verwendet. Bei der bevorzugten Ausführungsform, bei der die Bildmerkmalsinformation ein Merkmalsvektor ist, wird vorzugsweise eine euklidische Abstandsmeßik verwendet, um Dokumentenbilder zu finden, die dem Beispiel ähnlich sind. Alternativ kann ein Mahalanobis-Abstand, ein Hausdorff-Abstand oder ein gewichtetes Übereinstimmungsschema verwendet werden, um ähnliche Dokumentenbilder zu finden. Bei einer Ausführungsform nutzt das Suchverfahren den Vorteil einer Gruppierung bzw. einer Clusterbildung in der Bilddatenbank, um Clusterknoten zu finden, die Punkte darstellen, die in der Nähe des Beispielbildes in dem Bildmerkmal-Vektorraum liegen.

Im Schritt 208 werden ähnliche Dokumentenbilder dargestellt. Bei einer Ausführungsform werden die drei Bilder in der Datenbank, die am ähnlichsten sind, dargestellt. Alternativ werden Piktogramme bzw. Ikonen dargestellt, die Cluster bzw. Gruppen darstellen, die Bilder mit ähnlichen Charakteristiken enthalten.

Im Schritt 210 kann der Benutzer eines der dargestellten Dokumentenbilder oder Clusterpiktogramme als einen Schlüssel für die weitere Suche wählen. Das System nutzt dann das ausgewählte Bild oder das Clusterpiktogramm als ein neues Beispiel-Dokumentenbild und kehrt zum Schritt 206 zurück.

Bei einer Ausführungsform beinhaltet die Bilddatenbank Dokumentenbilder in verschlüsselter Form. Die Merkmalsinformation, die verwendet wird, um zu suchen, ist jedoch unverschlüsselt. Wenn gewünschte Dokumentenbilder identifiziert sind, werden sie für die Anzeige entschlüsselt. Gemäß dem Stand der Technik werden verschlüsselte Dokumentendatenbanken durchsucht, indem unverschlüsselte Schlüsselwort-Information verwendet wird. Der Lösungsansatz, der durch die vorliegende Erfindung bereitgestellt wird, ist dahingehend vorteilhaft, daß die unverschlüsselte Bildmerkmalsinformation nicht leicht verständlich ist und immer

noch eine Sicherheitsmaßnahme bereitstellt. Im Gegensatz dazu können die Schlüsselwörter, die gemäß dem Stand der Technik unverschlüsselt geblieben sind, leicht verstanden werden, wodurch die Sicherheit gefährdet ist.

Bei einer verwandten Ausführungsform beinhaltet die Bilddatenbank Dokumentenbilder in einer komprimierten Form, während die Merkmalsinformation, die für die Suche verwendet wird, nicht komprimiert ist.

10 Wenn gewünschte Dokumentenbilder identifiziert werden, werden sie zur Anzeige dekomprimiert.

Benutzerschnittstellen zum Entwickeln eines Beispielbildes

15 Fig. 3 zeigt einen Benutzerschnittstellenbildschirm 300 zum Entwickeln eines synthetischen Dokumentbildes, basierend auf Kategorieauslektionen als ein Beispiel, um eine Suche in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung zu starten. Ein Benutzerschnittstellenbildschirm 300 kann auf einem Touchscreen 40 realisiert werden, wobei der Benutzer Selektionen vornimmt, indem die Auswahl von Kategorien spezifiziert wird, indem die vielfältigen angezeigten 20 Tasten bzw. Schaltflächen berührt werden. Alternativ kann eine Maus oder eine andere vergleichbare Eingabevorrichtung verwendet werden, um die verschiedenen Kategorien auszuwählen.

Eine erste Spalte 302 von Schaltflächen, die den Titel 30 "Struktur" bzw. "Textur" trägt, ermöglicht es dem Benutzer einen Dokumententyp auszuwählen. Es werden Schaltflächen bzw. Tasten für "Zeitung" 304, "Geschäftsbrief" 306, "Journal/Magazin" 308, "Katalog/Broschüre" 310 und "Handgeschriebenes" 312 bereitgestellt. Der Benutzer wählt den Dokumententyp bezüglich der Dokumente die er/sie sucht.

Eine zweite Spalte 314 von Schaltflächen bzw. Tasten, die den Titel "Layout" trägt, erlaubt es dem Bediener, die Anzahl von Spalten von Text auszuwählen, die in den 40 Dokumenten gefunden werden, nach denen er/sie sucht. Schaltflächen werden für eine Spalte 316, für zwei Spalten 318 und für drei Spalten 320 bereitgestellt.

Eine dritte Spalte 322 von Schaltflächen bzw. Tasten, die den Titel "graphische Gestaltung" trägt, erlaubt es dem Benutzer die relative Anordnung eines Dokumententitels, von Text und eines graphischen Bildes auszuwählen, wie sie in einem gewünschten Suchergebnis gefunden werden würde. Jede Schaltfläche zeigt eine mögliche Anordnung. Eine Schaltfläche 324 zeigt einen Titel 50 326, der sich oberhalb eines graphischen Bildes 328 und eines Textes 330 befindet. Eine Schaltfläche 338 zeigt einen Titel 340 oberhalb eines graphischen Bildes 342. Eine Schaltfläche 344 zeigt ein graphisches Bild 346 in der Nähe des Textes 348.

55 Indem eine Auswahl aus jeder der drei Spalten getroffen wird, kann der Benutzer schnell das allgemeine Erscheinungsbild des gewünschten Dokuments oder der gewünschten Dokumente übermitteln. Nachdem der Benutzer seine bzw. ihre Selektionen gemacht hat, startet die Aktivierung einer "Finde"-Schaltfläche 350 eine Suche, die auf einer Bildmerkmalsinformation basiert, die den Selektionen des Benutzers entspricht. Diese Selektionen des Benutzers können als derartige verstanden werden, die ein Beispiel-Dokumentenbild spezifizieren, das als ein Schlüssel für die Suche in derselben Art und Weise dient, wie dies ein abgetastetes Bild tun würde. Natürlich sind die verschiedenen Kategorien und die Auswahl, die in Fig. 4 gezeigt ist, nur beispielhaft.

Fig. 4 zeigt eine vereinfachte Darstellung eines Benutzerschnittstellenbildschirms 400, bei dem ein Benutzer graphische Werkzeuge verwenden kann, um ein Beispiel-Dokumentbild zu entwickeln, um eine Suche in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung zu starten. Ein Benutzerschnittstellenbildschirm 400 bietet mehr Flexibilität bei der Spezifizierung eines Beispielbildes, als das System, das unter Bezugnahme auf Fig. 3 beschrieben ist.

Der Benutzer folgt einer Prozedur entlang der Rute, die er bei typischen computerunterstützten Zeichenpaketen folgen würde. Der Benutzer wählt einen Dokumentkomponententyp, wie zum Beispiel eine Kopfreile, Autor, Titel, Hauptkörper, Fußnote, Bild etc., indem eine Palette 402 verwendet wird. Eine Palette 404 wird verwendet, um eine Fontgröße bzw. eine Schriftzeichensatzgröße für Textkomponenten zu spezifizieren und um zwischen Strichzeichnungs- und Halbtontabellen für Bildkomponenten zu unterscheiden. Die Auswahl einer Dokumentkomponente kann die Verwendung einer Maus 38, um einen Zeiger über den relevanten Palettenabschnitt zu positionieren, das Niederdrücken einer Taste auf der Maus 38, das Positionieren des Zeigers bei einem gewünschten Punkt innerhalb eines Dokumentbildes 406 und das erneute Niederdrücken der Taste beinhalten. Die Größe und Position der Dokumentenkomponente kann dann justiert werden, indem Techniken, die für computerunterstützte Zeichenpakete typisch sind, verwendet werden.

Zahlreiche Variationen sind möglich bei dieser Benutzerschnittstelle. Um dem Benutzer eine Rückwirkung bzw. eine Systemreaktion bereitzustellen, kann das System einen aktuellen Text oder Grafikbilder in den bezeichneten Positionen innerhalb des Dokumentenbildes 406 darstellen. Ebenso kann der Benutzer die Dokumentenstruktur mit größerer Genauigkeit spezifizieren, indem die Fontgröße oder der Typenstil der Textbereiche des Dokumentenbildes spezifiziert wird.

Nach einer Reihe von Interaktionen bzw. Wechselwirkungen mit dem Benutzerschnittstellenbildschirm 400, ergibt sich ein Bild, wie zum Beispiel ein Dokumentenbild 406 und dieses Bild ist für eine Verwendung als ein Beispielbild für eine Suche verfügbar. Eine symbolische Darstellung 408 zeigt die Ergebnisse einer Suche, die auf einem Beispielbild 406 basiert.

In Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung wird eine Textabfrage mit einer Struktursuche kombiniert. Eine Textabfrage würde durch die Entwicklung eines Beispielbildes mittels irgendeiner Technik, die eine der oben beschriebenen beinhaltet, formuliert werden. Eine Textabfrage würde eine oder mehrere Text-Zeichenketten bzw. Text-Strings beinhalten, die bei einem gewünschten Dokument gefunden werden würden. Die Suchergebnisse würden auf Bilder beschränkt werden, die die Kriterien von sowohl der Textabfrage als auch der Textsuche erfüllen. Diese besondere Abfragetechnik ist natürlich auf Bilddatenbanken beschränkt, für die OCR-Ergebnisse verfügbar sind. Als erstes würden Dokumente mit Bildern wiedergefunden werden, die das gewünschte Muster an Text-Strings beinhalten. Die wiedergewonnenen Bilder werden dann nach Bildern durchsucht, die eine Struktur aufweisen, die der spezifizierten ähnlich ist.

Alternativ wird eine Bildstruktur verwendet, um die Ergebnisse zu organisieren, die durch eine Textabfrage wiedergegeben werden. Als erstes werden Dokumentenbilder, die das gewünschte Muster an Text-Strings beinhalten, wiedergewonnen. Cluster bzw. Gruppen

von Dokumenten mit ähnlichen Strukturen werden dann zusammen gruppiert. Diese Gruppen von Dokumenten werden dem Benutzer angezeigt, indem zum Beispiel eine Overlay-Technik verwendet wird, die früher beschrieben wurde. Diese Technik stellt einen einfachen Weg bereit, um zwischen Dokumenten von unterschiedlichen Quellen Unterscheidungen zu treffen. Man betrachtet zum Beispiel eine Datenbank, die Bilder von Zeichnungen und Artikeln wissenschaftlicher Journale beinhaltet. Eine Textabfrage kann Dokumente von beiden Typen von Quellen wiedergeben. Jedoch kann der Benutzer nur in den Ergebnissen von einer von ihnen interessiert sein. Die Bildstruktur-Clusterbildung stellt ein Verfahren für einen Benutzer bereit, um schnell Zeitungsartikel von Artikeln wissenschaftlicher Journale zu unterscheiden.

Suchergebnisanzeige

Fig. 5 zeigt, wie Suchergebnisse für eine geclusterte Bilddatenbank bzw. eine gruppierte Bilddatenbank in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung dargestellt werden kann. Eine Bilddatenbank mit Clustern ist eine, wo Bilder in einer hierarchischen Struktur gemäß ihrer Position in dem Raum zusammen gruppiert werden, der durch den Bildmerkmalsvektor festgelegt ist. Die Datenbasis wird in eine oder mehrere Gruppen aufgeteilt. Jede Gruppe kann wiederum aus einer oder mehreren Untergruppen bestehen. Eine Beschreibung der Clusterbildung kann in: Duda & Hart, "Pattern Classification & Scene Analysis" (John Wiley & Sons 1973), deren Inhalt hiermit durch Bezugnahme aufgenommen wird.

Falls es ein einziges Hierarchieniveau gibt, werden bei einer Abfrage alle Bilder in dem Cluster, die dem Beispieldokument am nächsten sind, in dem Bildmerkmals-Vektorraum zurückgegeben. Fig. 5 zeigt die Ergebnisse einer Durchblätterungssuche, die in dem Kontext bzw. im Zusammenhang einer Vielfach-Niveau-Hierarchie von Clustern dargestellt sind. Das Niveau 500 zeigt Piktogramme, die Clusterknoten darstellen. Der Benutzer steigt in der Hierarchie weiter ab, indem eines der Piktogramme als ein Schlüssel für die weitere Suche ausgewählt wird. Schließlich kann der Benutzer ein Blatt-Cluster ("leaf cluster") erreichen, d. h. eine Gruppe von tatsächlichen Bildern im Gegensatz zu Cluster-Knoten. Diese Bilder werden dann dargestellt.

Ein Weg der Realisierung der auf Cluster basierenden Suche ist es, die Clusterstruktur jedesmal neu zu berechnen, wenn eine Abfrage vorgelegt bzw. gefordert wird. Eine andere Art und Weise der Realisierung der auf Cluster basierenden Suche ist es, die Clusterstruktur vor der Suche zu berechnen, indem vorzugsweise Strukturinformation verwendet wird. Der zuvor beschriebene Suchprozeß gleicht dann im wesentlichen der Navigation durch einen Baum.

Fig. 6 zeigt eine verbesserte Anzeige für Dokumente mit mehreren Seiten in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung. Häufig stellt ein gewünschtes Dokumentbild eine Seite einer Bilddatenbank dar. Zum Beispiel sind in der gutbekannten Berkeley-Bilddatenbank Reports bzw. Berichte von 50 Seiten üblich. Viele Seiten dieser Reports sind nur Text und neigen dazu, ähnliche Strukturmerkmale aufzuzeigen, was ein Wiederfinden schwierig macht. Die Seiten, die die interessantesten Merkmale für den Benutzer für das Durchblättern bzw. Durchstöbern bereitstellen, sind die Titelseite, die Seite mit der Inhaltsanga-

be, Diagramme, Figuren etc. Diese stellen ebenso die Seiten dar, die die besten Merkmale bereitstellen, von denen aus weitere Abfragen der Datenbank vorgenommen werden können.

Vorzugsweise erscheint ein Dokument mit mehreren Seiten zu Beginn als ein einziges Bild, wenn ein Suchergebnis dargestellt wird. Jedoch kann der Benutzer bewirken, daß das gesamte Dokument erscheint, indem dieses Bild ausgewählt wird. Eine Anzeige 602 zeigt die Suchergebnisse einschließlich der mehreren Seiten 606 desselben Reports, die ein ähnliches Erscheinungsbild aufweisen. Man kann erkennen, daß für Reports mit Hunderten von Seiten die Anzeige bald mit sich wiederholenden Bildern überhäuft ist. Eine verbesserte Darstellung 604 zeigt mehrere Seiten 606 desselben Reports, wobei dessen Bilder mit einem leichten Versatz von Seite zu Seite überlagert sind. Eine verbesserte Anzeige 604 zeigt eine erheblich geringere Überhäufung.

Um die verbesserte Darstellung zu erreichen, wenn ein Report mit mehreren Seiten dargestellt werden soll, werden Seiten mit einem im wesentlichen gleichen Erscheinungsbild identifiziert, indem ihre Seitenmerkmalsinformation verglichen wird.

Die Bilder dieser Seiten werden dann dargestellt, wobei ihre Bilder mit einem Offset von Seite zu Seite überlagert werden.

Merkmalvektorextraktion

Fig. 7A zeigt die Elemente eines Merkmalvektors 700, die von einem Dokumentbild in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung extrahiert werden. Vorzugsweise beinhaltet ein Merkmalvektor 700 80 Elemente, die vier Gruppen von 20 Elementen beinhalten. Eine erste Gruppe von 702 von 20 Elementen stellt ein Histogramm von verbundenen Komponentengrößen dar und wird weiter unter Bezugnahme auf Fig. 7B diskutiert. Eine zweite Gruppe 704 von 20 Elementen stellt eine Verteilung von Interessenpunkten durchgehend über die Zellen des Dokumentbildes dar und wird weiter unter Bezugnahme auf Fig. 7C diskutiert. Eine dritte Gruppe 706 von 20 Elementen stellt ein Vertikalprojektions-Histogramm einer Verbindungskomponentendichte dar und wird weiter unter Bezugnahme auf Fig. 7D diskutiert. Eine vierte Gruppe 708 von 20 Elementen stellt eine Verteilung von Verbindungskomponenten bzw. verbundenen Komponenten durchgehend über Zellen des Dokuments dar und wird weiter unter Bezugnahme auf Fig. 7E diskutiert.

Fig. 7B zeigt ein Verbindungskomponenten-Größenhistogramm 710, wie es in einem Merkmalvektor 702 in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfindung inkorporiert werden würde. Das Histogramm 710 basiert auf einer Verbindungskomponenten-Größeninformation, die von dem Bild erhalten wird. Eine einzige Verbindungskomponente stellt einen Satz von benachbarten schwarzen Pixeln in einem Bild dar. Um die Größe einer Verbindungskomponente zu bestimmen, findet die bevorzugte Ausführungsform das Gebiet der rechteckigen Box minimaler Größe, das sie umschließt. Alle anderen Arten der Bestimmung einer Verbindungskomponentengröße gehören ebenfalls zum Umfang der vorliegenden Erfindung. Zum Beispiel kann die diagonale Abmessung der umfassenden bzw. einschließenden rechteckigen Box die Verbindungskomponentengröße sein. Schließlich könnte noch, wenn dies zusätzlich erwünscht wird, mit dem Bild eine Filterung durchgeführt werden, bevor eine Verbindungskompo-

nentengröße bestimmt wird.

Horizontale und vertikale Achsen 712 und 714 des Histogramms werden vorzugsweise zu einer logarithmischen Skala normalisiert. Weiter könnte man zusätzlich noch eine Gewichtung mit dem Histogrammbins vornehmen. Ein Plot 716 zeigt das Histogramm für eine symbolische Seite bzw. repräsentative Seite eines Magazins. Ein Plot 718 zeigt das Histogramm für eine symbolische bzw. repräsentative Seite eines technischen Journals. Wie man sehen kann, ist die Information von diesem Histogramm nützlich, um zwischen verschiedenen Typen von Dokumenten zu unterscheiden.

Die zweite Gruppe 704 von 20 Merkmalvektorelementen basiert auf einer Interessenpunktinformation. 15 Interessenpunkte sind Pixel innerhalb eines Bildes, die innerhalb von Gebieten liegen, die unterschiedliche bzw. zu unterscheidende Merkmale aufweisen. Jeder Interessenpunkt weist einen identifizierten bzw. erkannten Interessenpegel auf. In der Fachwelt ist ebenso bekannt, daß Interessenpunkte ("interest points") und ihre zugeordneten Interessenpegel ("interest levels") gefunden werden, indem sogenannte Interessenoperatoren auf die Pixel eines Bildes angewendet werden. Die bevorzugte Ausführungsform sieht einen Vorteil aus einem modifizierten Moravec-Operator, der dem ähnelt, der in der US-Patentanmeldung mit der Seriennummer 08/527,286 "SIMULTANEOUS REGISTRATION OF MULTIPLE IMAGE FRAGMENTS" beschrieben ist, die auf den Anmelder dieser Anmeldung überschrieben ist und deren Inhalt hiermit durch Bezugnahme mit aufgenommen wird.

Für jeden Pixel wird der modifizierte Moravec-Varianzoperator an jedem der sechs Räden von dem Pixel aus bestimmt. Für einen gegebenen Pixel und einen Radius stellt der Interessenpegelwert bzw. Interessenniveauwerte die Varianz der Pixelwerte der vier Pixel dar, die sich um einen gegebenen Radius weg, oberhalb, unterhalb, links und rechts von dem gegebenen Pixel befinden. Typischerweise sind Dokumentenbilder binär und die Pixelwerte sind entweder Null oder stellen einen Skalenendwert bzw. einen Höchstwert dar. Die Interessenpegelwerte werden in drei Bits sortiert, in einen Niedriginteressenwert, einen mittleren Interessenwert und in einen Hochinteressenwert. Andere Interessenoperatoren können ebenso verwendet werden. Einen Überblick über Interessenoperatoren wird in Yan "Interest Operators and Fast Implementation" International Archives of Photogrammetry and Remote Sensing, Band 27-H, Japan, Seiten 491–500, 1988, gefunden, deren Inhalt hiermit unter Bezugnahme uneingeschränkt mit aufgenommen wird.

Jede Kombination aus Radius und Interessenpegelbin wird einem Element in einer zweiten Gruppe 704 zugewiesen. Der Wert des Elements stellt eine normalisierte Darstellung der Anzahl von Pixel mit dem Interessenpegelwert und Radius, der dem Element zugeordnet ist, dar. Da es sechs Räden gibt und drei Interessenpegel-Wertebins, macht dies 18 von 20 Elementen in der zweiten Gruppe 704 aus.

Die zwei verbleibenden Elemente der zweiten Gruppe 704 stellen spezielle Werte dar. Ein Element beinhaltet einen Wert, der dem Verhältnis der Anzahl von schwarzen Pixel in dem Bild zu der gesamten Anzahl von Pixel entspricht. Ein anderes Element hält einen Wert, der einer gesamten Anzahl von Schwarz-zu-Weiß-Übergängen entspricht, die bei einer horizontalen Abtastung durch das Bild entdeckt werden.

Fig. 7C stellt die Nützlichkeit einer Interessenpunkt-

information für die Unterscheidung von Dokumentenbildern dar. Fig. 7C zeigt einen Buchstaben "t" 720, wobei die Interessenpegel bei verschiedenen Teilen des Buchstabens vorhanden sind. Ein Pixel 722 an einer Ecke weist einen hohen Interessenpegel auf. Ein Pixel 724 im Inneren weist einen niedrigen Interessenpegel auf. Ein Pixel 726 entlang einer Kante weist einen mittleren Interessenpegel auf.

Fig. 7D zeigt ein Vertikalprojektion-Verbindungs-komponentenhistogramm 728, wie es innerhalb eines Merkmalvektors in Übereinstimmung mit einer Ausführungsform der vorliegenden Erfahrung inkorporiert werden würde. Um die dritte Gruppe 706 des Merkmalvektors zu erhalten, wird das Bild in 20 Vertikalstreifen aufgeteilt. Jedes Element der dritten Gruppe entspricht einer Population bzw. einer Gesamtheit von Verbindungskomponenten mit einer Größe oberhalb einer minimalen Schwelle in einem bestimmten Vertikalstreifen. Bei einer bevorzugten Ausführungsform stellt dieses Minimum eine Höhe und eine Breite dar, die beide größer als ein Pixel sind. Diese dritte Gruppe 706 ist besonders bei der Unterscheidung zwischen Dokumenten nützlich, die eine unterschiedliche Anzahl von Spalten im Text aufweisen. Ein erstes Dokument 730 mit einer einzigen Textspalte ergibt die Projektions-Histogrammdarstellung 732. Ein zweites Dokument 734 mit zwei Textspalten ergibt ein Projektionshistogramm 736. Wie man sehen kann, weist das Projektionshistogramm 732 einen Peak auf, der der einzigen Spalte entspricht, wohingegen das Projektionshistogramm 736 zwei Peaks aufweist, die den zwei Spalten entsprechen.

Fig. 7E zeigt, wie die Verteilung von Verbindungs-komponenten zwischen Gitterzellen, wie sie innerhalb eines Merkmalsvektors in Übereinstimmung mit seiner Ausführungsform der Erfahrung inkorporiert werden würden, arbeitet, um eine Unterscheidung zwischen Dokumenten mit unterschiedlichen Anordnungen von Komponenten zu treffen. Die vierte Gruppe 708 von Elementen wird erhalten, indem das Bild in ein 5×4 -Feld 738 von Zellen aufgeteilt wird und indem die Population bzw. Gesamtheit von Verbindungskomponenten oberhalb einer Schwellengröße in jeder Zelle gefunden wird. Bei der bevorzugten Ausführungsform stellt diese Komponentengröße eine Höhe und eine Breite dar, die größer als ein Pixel ist. Die Population von Verbindungskomponenten bzw. verbundenen Komponenten einer jeden Zelle wird dann in Bezug auf eine Gesamtanzahl von Verbindungskomponenten für das Bild als Ganzes normalisiert. Diese vierte Gruppe 708 ist besonders bei der Unterscheidung zwischen Dokumenten, wie zum Beispiel Dokumente 740 und 742 nützlich, die unterschiedliche Anordnungen von Dokumentkomponenten bzw. Dokumenten-Komponenten aufweisen.

Die Extraktion von Bildmerkmals-Information aus Beispiel-Dokumentenbildern, die durch Kategoriselektion entwickelt wurden, wie unter Bezugnahme auf Fig. 3 beschrieben wurde, oder die durch eine graphische Benutzerschnittstelle entwickelt wurden, wie unter Bezugnahme auf Fig. 4 beschrieben wurde, folgt einem modifizierten Verfahren bzw. einer modifizierten Prozedur. Bei dem Kategoriselektionsschema von Fig. 3 gibt es eine beschränkte Anzahl von möglichen Beispiels-Dokumentenbildern. Ein Verfahren besteht darin, für jede Kombination von Kategoriselektionen ein entsprechendes aktuelles Dokumenten-Bild mit den gewählten Charakteristiken bereitzustellen. Dieses Bild kann dann als die Basis für eine Bildmerkmalsextraktion

dienen, wie sie oben unter Bezugnahme auf die Fig. 7A – 7E beschrieben wurde.

Für Beispielbilder, die entwickelt wurden, indem die Techniken verwendet werden, die unter Bezugnahme auf Fig. 4 diskutiert wurden, besteht ein Verfahren darin, ein aktuelles Dokumenten-Bild zu konstruieren, indem Beispieldatamaterial bzw. Abtastdatamaterial kombiniert wird, das jeder ausgewählten Komponente entspricht. Bezuglich Textkomponenten wird tatsächlicher Text mit aufgenommen. Bezuglich Graphiken wird ein Abtastgraphikbild bzw. ein Beispielgraphikbild mit aufgenommen bzw. mit eingeschlossen. Das so konstruierte aktuelle Dokumentenbild dient als Grundlage für eine Bildmerkmalsextraktion.

Suchergebnisse

Fig. 8 zeigt ein repräsentatives Beispiel-Dokumentenbild 802, das als ein Beispiel für eine Suche in Übereinstimmung mit der vorliegenden Erfahrung verwendet wird, und Suchergebnisse. In dem ein Beispieldbild 802 als der Schlüssel für eine Suche verwendet wird, werden Bilder 804 und 806 und 808 von einer Datenbank von 971 Bildern einschließlich Journalen, Briefen, Magazinen, Broschüren bzw. Katalogen, Zeitungen, Katalogen und handgeschriebenen Text. Das Beispieldbild 802 stellt einen Geschäftsbrief dar und alle Suchergebnisse sind ebenfalls Geschäftsbriefe. Wie man sehen kann, stellt das Suchsystem der Erfahrung Bilder bereit, die vom Anblick her dem Beispieldbild ähneln.

Gemäß der Erfahrung können zum Beispiel die Techniken der vorliegenden Erfahrung ebenso auf Grauskalenbilder angewendet werden, obwohl die obige Beschreibung auf binäre Bilder bezugnimmt. Interessenpunkt-Techniken sind in der Fachwelt für Grauskalenbilder gut bekannt. Verbindungskomponenten-Information kann von einem Grauskalenbild erhalten werden, indem zuerst eine Schwellenfunktion auf die Grauskalenwerte angewendet wird und dann benachbarte Gruppen von Pixeln mit Werten oberhalb der gewählten Schwelle gefunden werden.

Die Suchtechniken, die oben beschrieben wurden, können ebenso auf Farbbilder mit passenden Änderungen des Bildmerkmalsvektors angewendet werden. Zum Beispiel kann der Bildvektor derartig ausgedehnt werden, daß er Elemente enthält, die den Median-CMYK-Werten über verschiedene bzw. vielfältige Zellen des Bildes entsprechen.

Ebenso kann für eine verbesserte Genauigkeit beim Vergleich eines visuellen Erscheinungsbildes von Bildern ein Merkmalsvektor mit einer viel größeren Anzahl von Elementen verwendet werden. Zum Beispiel kann das Dokumentenbild in 20 Zellen aufgeteilt werden. Ein 60-Elementvektor, der den ersten drei Gruppen des Merkmal-Elementvektors 700 entspricht, kann dann für jede Zelle berechnet werden. Die Ansammlung aller dieser 60-Elementvektoren würde einen 1.200-Element-Merkmalvektor darstellen, der als Basis zum Durchsuchen und Durchblättern einer Bilddatenbank dienen kann.

Erfahrungsgemäß kann z. B. das offenbare Verfahren auch über größere Entfernung durchgeführt werden. So kann nach dem Schritt des Erhaltens der Bildmerkmalsextraktion (siehe z. B. Anspruch 1), diese Information an die Datenbank übermittelt werden, die sich z. B. an eine entfernt gelegene bzw. an einem anderen Ort, als den Beispieldatenerzeugungsort bzw. Bildmerkmalsextraktionsort befindet. Die Datenbank sen-

det dann Suchergebnisse zur Weiterverarbeitung zurück. Aus diese Art und Weise müssen nicht sämtliche zu einem Thema gehörigen Dokumente von der Datenbank zur Durchsicht übermittelt werden, da eine gezielte Vorauswahl getroffen wurde. Dadurch wird die Datenübertragungszeit entscheidend verkürzt und der Speicherbedarf am Abfragegerät bzw. am Abfrageende des Geräts sinkt.

Weiter kann z. B. durch die Bereitstellung von den einzelnen Datenbank-Bilddokumenten zugeordneten Merkmalsvektoren, die Suche erheblich beschleunigt werden, wobei die einzelnen Merkmalsvektoren z. B. als eine Art komprimierte, wenn auch verlustbehaftete Darstellung der Bilddokumente aufgefaßt werden können. So kann z. B. die Übermittlung von Merkmalsvektoren in Kombination mit Datenbanken an einem oder an beiden Enden des Übermittlungsweges als Kompressions-/Dekompressionsverfahren zur Übermittlung von Bildern aufgefaßt werden.

Die Erfindung läßt sich beispielsweise wie folgt zusammenfassen:

Verfahren und Apparat zum Abfragen einer Dokument-Bilddatenbank, basierend auf einer Struktur, d. h. auf analytisch unterscheidbaren Mustern in den Dokument-Bildern der Datenbank. Gemäß der Erfindung kann eine Dokumentbilddatenbank für Dokumente mit einer bestimmten Struktur auf eine Vielfalt von Weisen durchblättert werden. Zum Beispiel kann ein Benutzer ein Beispiel-Dokumentbild mit einem Erscheinungsbild, das dem gewünschten Dokument ähnelt, eingeben. Alternativ kann der Benutzer eine einfache Schnittstelle verwenden, um ein synthetisches Dokument bzw. künstliches Dokument, basierend auf einer Auswahl weniger Kategorien festzulegen. Das synthetische Dokument würde dann als ein Beispiel für eine Suche dienen. Oder der Benutzer kann eine graphische Schnittstelle verwenden, um genauer ein Beispiel für eine Suche festzulegen. Somit liefert die Kenntnis des Benutzers von dem allgemeinen Erscheinungsbild des gewünschten Dokuments oder der Dokumente die Basis für die Suche.

Patentansprüche

1. Verfahren zum Durchsuchen einer Dokument-Bilddatenbank, das die folgenden Schritte aufweist: ein Beispiel-Dokumentbild wird entwickelt; eine Bildmerkmalsinformation über das Beispiel-Dokumentbild wird erhalten; die Dokument-Bilddatenbank wird durchsucht, indem die Bildmerkmalsinformation verwendet wird;

Suchergebnisse des Suchschrittes werden angezeigt; und eine Benutzereingabe wird angenommen bzw. angezeigt, die ein gegebenes Suchergebnis als einen Schlüssel für eine weitere Suche auswählt.

2. Verfahren nach Anspruch 1, bei welchem der Entwicklungsschritt ein Einscannen des Beispiel-Dokumentbildes umfaßt.

3. Verfahren nach Anspruch 1, bei welchem der Entwicklungsschritt die Annahme einer Benutzer-eingabe umfaßt, die Charakteristiken des Beispiel-Dokumentbildes spezifiziert.

4. Verfahren nach Anspruch 1, bei welchem der Entwicklungsschritt den Schritt umfaßt, wonach ein graphisches System betrieben wird, bei dem ein Benutzer Komponenten des Beispiel-Dokumentbildes spezifiziert und anordnet.

5. Verfahren nach Anspruch 1, bei welchem die

Bildmerkmalsinformation eine Information über Interessenpunkte in dem Beispiel-Dokumentbild umfaßt.

6. Verfahren nach Anspruch 1, bei welchem die Bildmerkmalsinformation eine Information über verbundene Komponenten in dem Beispiel-Dokumentbild umfaßt.

7. Verfahren nach Anspruch 1, bei welchem die Bildmerkmalsinformation Information über eine Verteilung von verbundenen Komponenten bzw. Verbindungskomponenten in dem Beispiel-Dokumentbild umfaßt.

8. Verfahren nach Anspruch 1, bei welchem der Schritt des Erhaltens den Schritt aufweist, wonach ein Bildmerkmalsvektor für das Beispiel-Dokumentbild bestimmt wird.

9. Verfahren nach Anspruch 8, bei welchem der Suchschritt den Schritt aufweist, wonach der Bildmerkmalsvektor für das Beispiel-Dokumentbild mit den Bildmerkmalsvektoren von Dokumenten-Bildern in der Datenbank verglichen werden.

10. Verfahren nach Anspruch 9, das weiter den Schritt aufweist, wonach eine Benutzereingabe angenommen wird, die einen Suchtextstring spezifiziert und wobei der Suchschritt weiter den Schritt aufweist, wonach der Suchtextstring mit OCR-Daten verglichen wird, die den Dokumenten-Bildern in der Datenbank zugeordnet sind.

11. Verfahren nach Anspruch 1, bei welchem der Schritt des Anzeigens das Anzeigen von einen oder mehreren Dokument-Bildern umfaßt, die strukturell dem Beispiel-Dokumentbild ähneln.

12. Verfahren nach Anspruch 1, bei welchem der Schritt des Anzeigens weiter das Anzeigen von Piktogrammen bzw. Ikonen als ein Suchergebnis umfaßt, die Cluster oder Gruppen von Bildern darstellen, die dem Beispiel-Dokumentbild ähneln.

13. Verfahren nach Anspruch 1, bei welchem die Dokument-Bilddatenbank Dokumenten-Bilder in verschlüsselter Form und Bildmerkmalsinformation für die verschlüsselten Dokumenten-Bilder aufweist.

14. Verfahren nach Anspruch 13, bei welchem das Suchen das Suchen der Bildmerkmalsinformation aufweist, die in der Dokument-Bilddatenbank gespeichert ist und bei welchem das Anzeigen ein Entschlüsseln von Suchergebnissen des Suchschrittes umfaßt.

15. Verfahren nach Anspruch 1, bei welchem die Dokument-Bilddatenbank Dokument-Bilder in komprimierter Form und Bildmerkmalsinformation für die komprimierten Dokumenten-Bilder aufweist.

16. Verfahren nach Anspruch 15, bei welchem das Suchen das Suchen der Bildmerkmalsinformation aufweist, die in der Dokument-Bilddatenbank gespeichert ist und bei welchem das Anzeigen ein Dekomprimieren von Suchergebnissen des Suchschrittes umfaßt.

17. Verfahren zum Extrahieren eines Merkmalsvektors aus einem elektronisch gespeicherten Bild, das die folgenden Schritte aufweist:

die Größe von Verbindungskomponenten bzw. verbundenen Komponenten wird über das ganze Bild gemessen;

Interessenpegel von Pixeln werden über das ganze Bild identifiziert;

die verbundenen Komponente bzw. Verbindungs-

komponenten und ihre Größen werden über das ganze Bild identifiziert;

ein Histogramm von verbundenen Komponentengrößen bzw. Verbindungskomponentengrößen werden als eine erste Gruppe von Elementen des Merkmalvektors gespeichert; und

ein Histogramm von Interessenpegeln von Pixeln wird als eine zweite Gruppe von Elementen des Merkmalvektors gespeichert.

18. Verfahren nach Anspruch 17, das weiter folgendes aufweist: ein Histogramm einer verbundenen Komponentenpopulation bzw. einer Verbindungs-komponentenpopulation in Vertikalabschnitten des Bildes wird als eine dritte Gruppe von Elementen des Merkmalvektors gespeichert.

19. Verfahren nach Anspruch 17, das weiter folgendes aufweist:

die Anzahl von verbundenen Komponenten bzw. Verbindungskomponenten werden in einer Vielzahl bzw. Anzahl von Regionen bzw. Bereichen des Bildes als eine vierte Gruppe von Elementen des Merkmalvektors gespeichert.

20. Verfahren zum Suchen nach einem Dokument in einer Dokument-Bilddatenbank, das die folgenden Schritte aufweist:

eine Benutzereingabe wird angenommen, die eine Kategorie des Dokuments spezifiziert;

eine Benutzereingabe wird angenommen, die eine Anzahl von Textspalten des Dokuments spezifiziert;

eine Benutzereingabe wird angenommen, die eine Organisation bzw. Aufbau des Dokuments spezifiziert;

ein künstliches Dokument bzw. synthetisches Dokument der Kategorie wird entwickelt, das die Anzahl von Spalten und die Organisation bzw. der Aufbau aufweist; und

die Dokument-Bilddatenbank wird nach Dokumenten durchsucht, die dem synthetischen bzw. künstlichen Dokument ähneln.

21. Verfahren nach Anspruch 20, bei welchem der Suchschritt weiter folgendes aufweist:

es wird nach Dokumenten gesucht, die eine Struktur aufweisen, die dem synthetischen Dokument bzw. künstlichen Dokument ähnelt.

22. Dokumentenbild-Durchstöberungssystem, das folgendes aufweist:

eine elektronische Speichereinheit, die eine Dokument-Bilddatenbank speichert;

eine Anzeige, die Dokumenten-Bilder anzeigt;

eine Prozessoreinheit bzw. Verarbeitungseinheit, die mit der elektronischen Speichervorrichtung und der Anzeige gekoppelt ist bzw. verbunden ist, wobei die Verarbeitungseinheit funktionsfähig ist, um:

ein Beispiel-Dokumentenbild zu entwickeln;

eine Bildmerkmalsinformation über das Beispiel-Dokumentenbild zu erhalten; und

die Dokumenten-Bilddatenbank nach Dokument-Bildern zu durchsuchen, die dem Beispiel-Dokumentenbild ähneln, wobei die Bildmerkmalsinformation verwendet wird.

23. Software auf einem Speichermedium, die folgendes aufweist:

Software zum Entwickeln eines Beispiel-Dokumentbildes;

Software zum Erhalten von Bildmerkmalsinformation über das Beispiel-Dokumentenbild;

Software zum Durchsuchen der Dokument-Bilddatenbank, indem die Bildmerkmalsinformation verwendet wird;

Software zum Anzeigen von Suchergebnissen der Suchsoftware; und

Software zum Annehmen einer Benutzereingabe, die ein angezeigtes Suchergebnis als einen Schlüssel für eine weitere Suche auswählt.

24. Software nach Anspruch 23, bei welcher die Entwicklungssoftware eine Software zum Annehmen einer Benutzereingabe aufweist, die Charakteristiken des Beispiel-Dokumentbildes spezifiziert.

25. Software nach Anspruch 23, bei welchem die Entwicklungssoftware weiter die Software zum Arbeiten eines graphischen Systems umfaßt, wobei ein Benutzer Komponenten des Beispiel-Dokumentbildes spezifiziert und anordnet.

26. Software nach Anspruch 23, bei welchem die Software zum Erhalten eine Software zum Extrahieren eines Bildmerkmalsvektors für das Beispiel-Dokumentenbild umfaßt.

27. Software nach Anspruch 26, bei welchem die Suchsoftware weiter Software zum Vergleichen des Bildmerkmalsvektors für das Beispiel-Dokumentenbild mit den Bildmerkmalsvektoren von Dokument-Bildern in der Datenbank aufweist.

28. Software nach Anspruch 27, die weiter Software zum Annehmen einer Benutzereingabe aufweist, die einen Suchtextstring spezifiziert und wobei die Suchsoftware weiter eine Software zum Vergleichen des Suchtextstrings mit OCR-Daten aufweist, die den Dokumentenbildern in der Datenbank zugeordnet sind.

29. Verfahren zum Anzeigen von Bildern von Dokumenten mit mehreren Seiten, das folgende Schritte aufweist:

Bildmerkmalsinformation von Bildern von Seiten von Dokumenten wird verglichen;

Seiten mit einem ähnlichen Erscheinungsbild werden basierend auf der Bildmerkmalsinformation identifiziert; und Seiten mit einem ähnlichen Erscheinungsbild werden übereinander mit einem Versatz zwischen ihnen überlagert angezeigt.

Hierzu 12 Seite(n) Zeichnungen

- Leerseite -

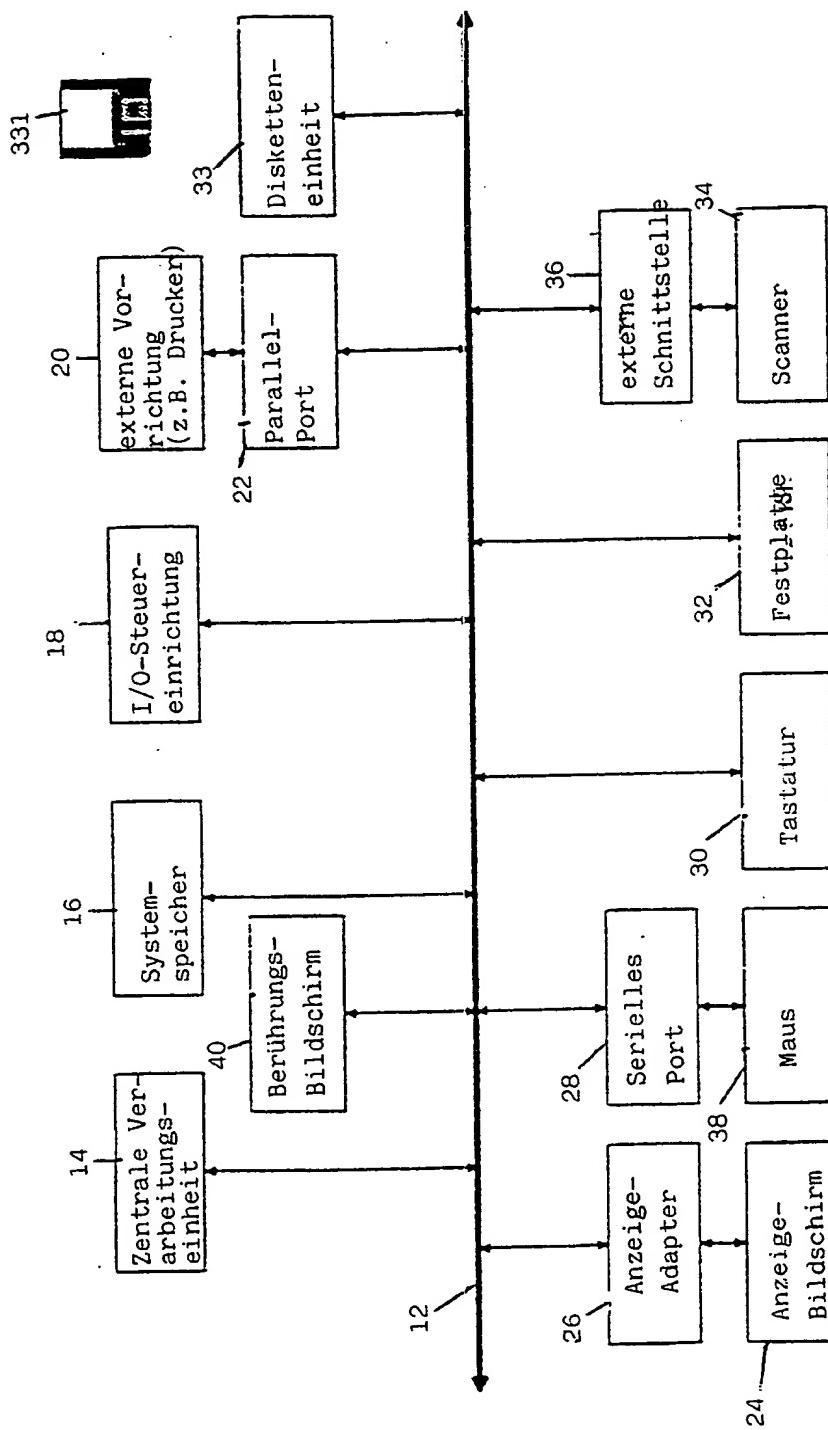


Fig. 1

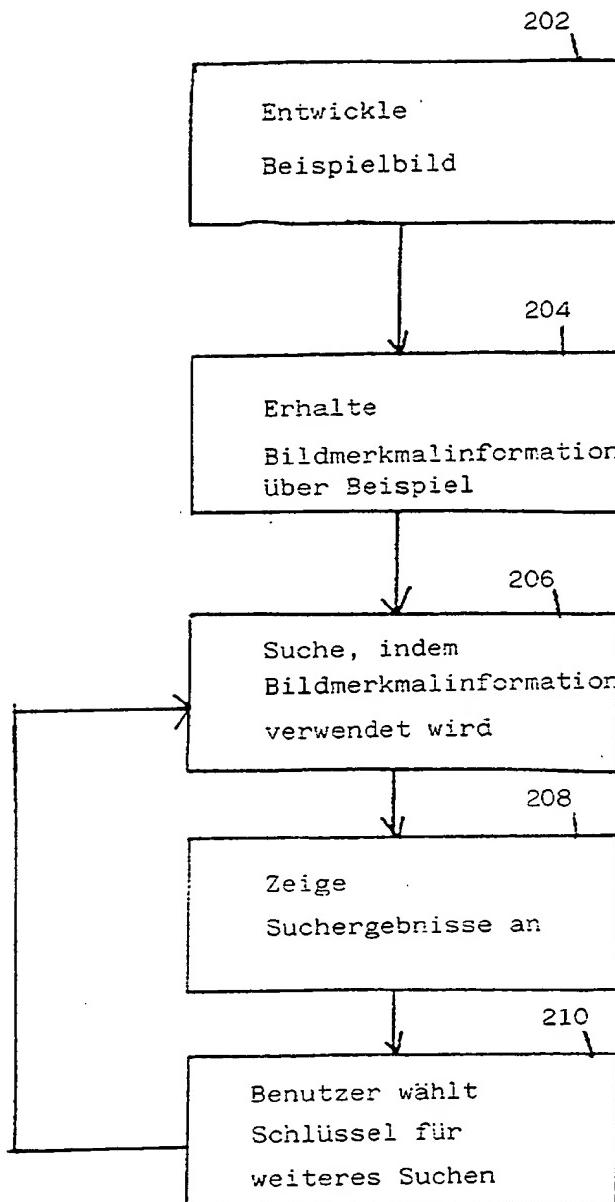


Fig. 2

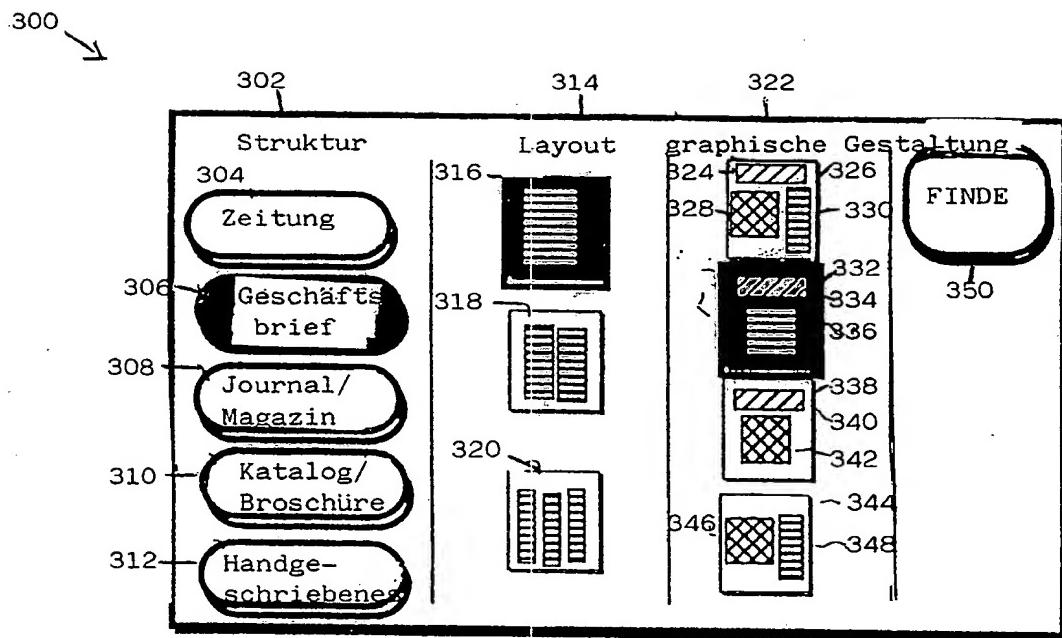
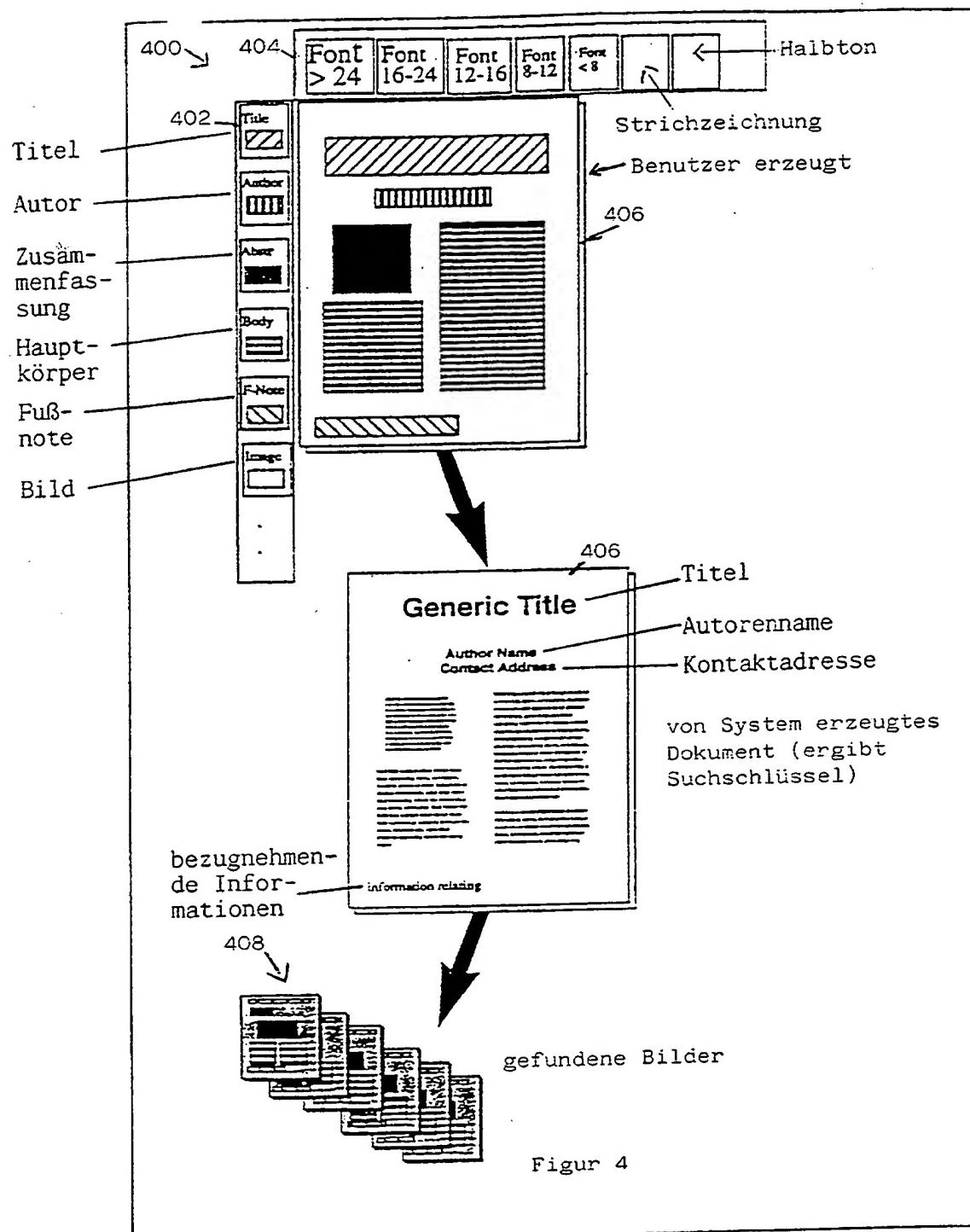


Fig. 3



Hierarchisches Durchstöbern

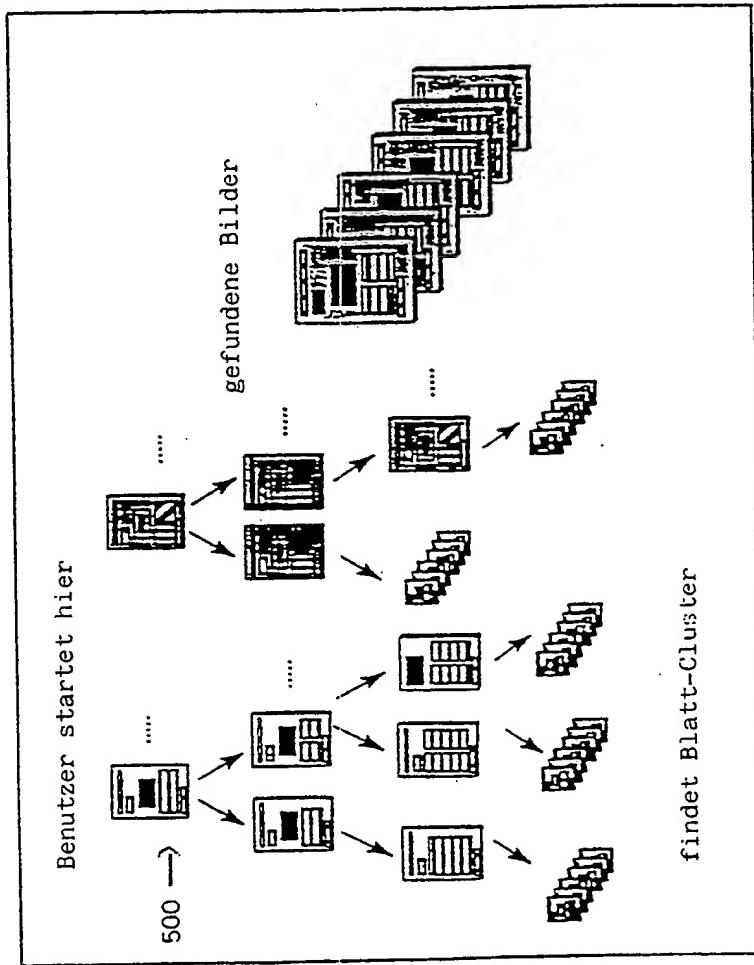


Fig. 5

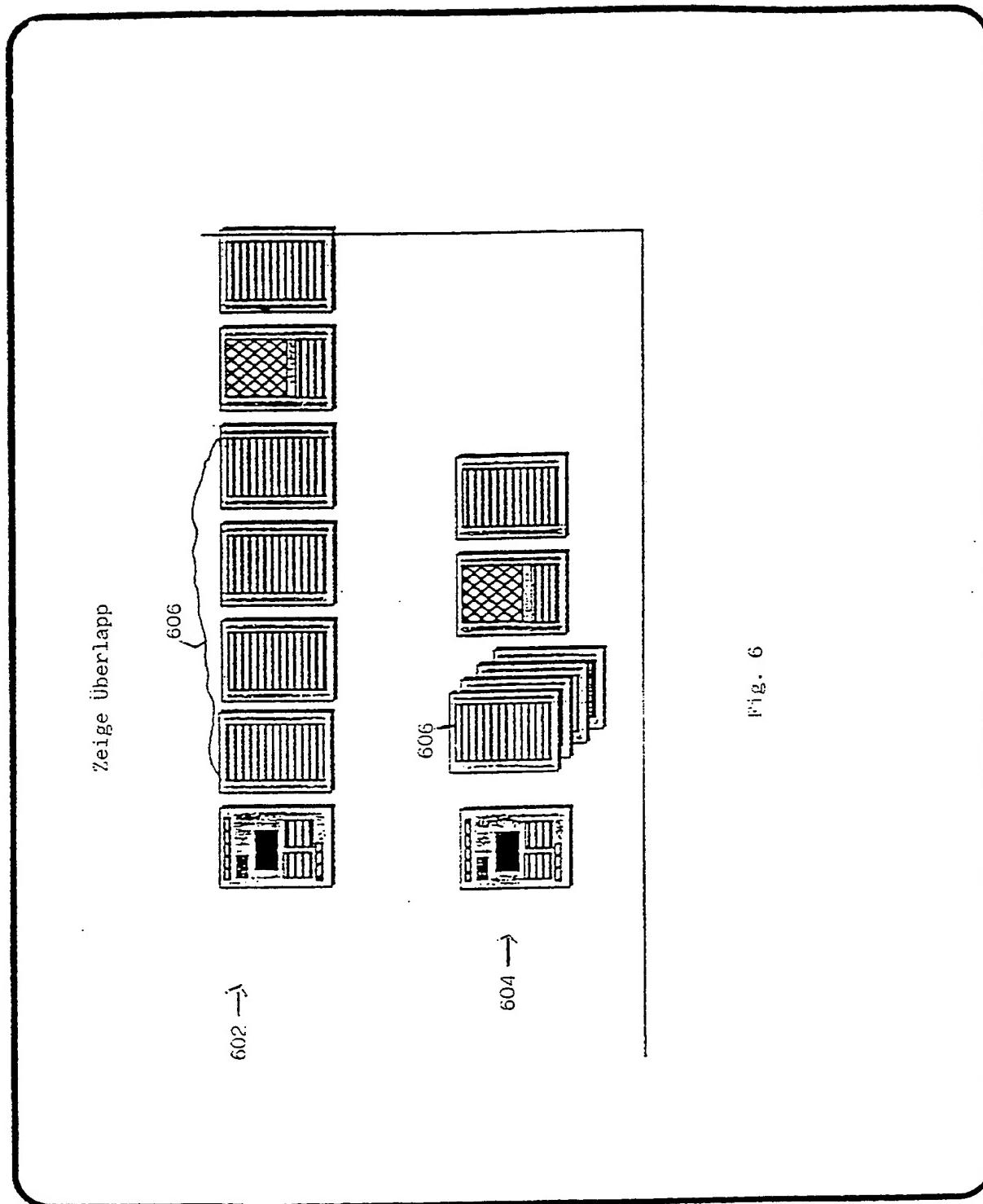


Fig. 6

Nummer:

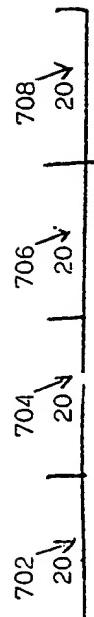
DE 197 08 265 A1

Int. Cl.⁶:

G 06 F 17/30

Offenlegungstag:

4. September 1997



700 ↗

Fig. 7A

702 036/516

Druck eines Histogramms von log log (Gebiet) von
Verbindungs komponenten

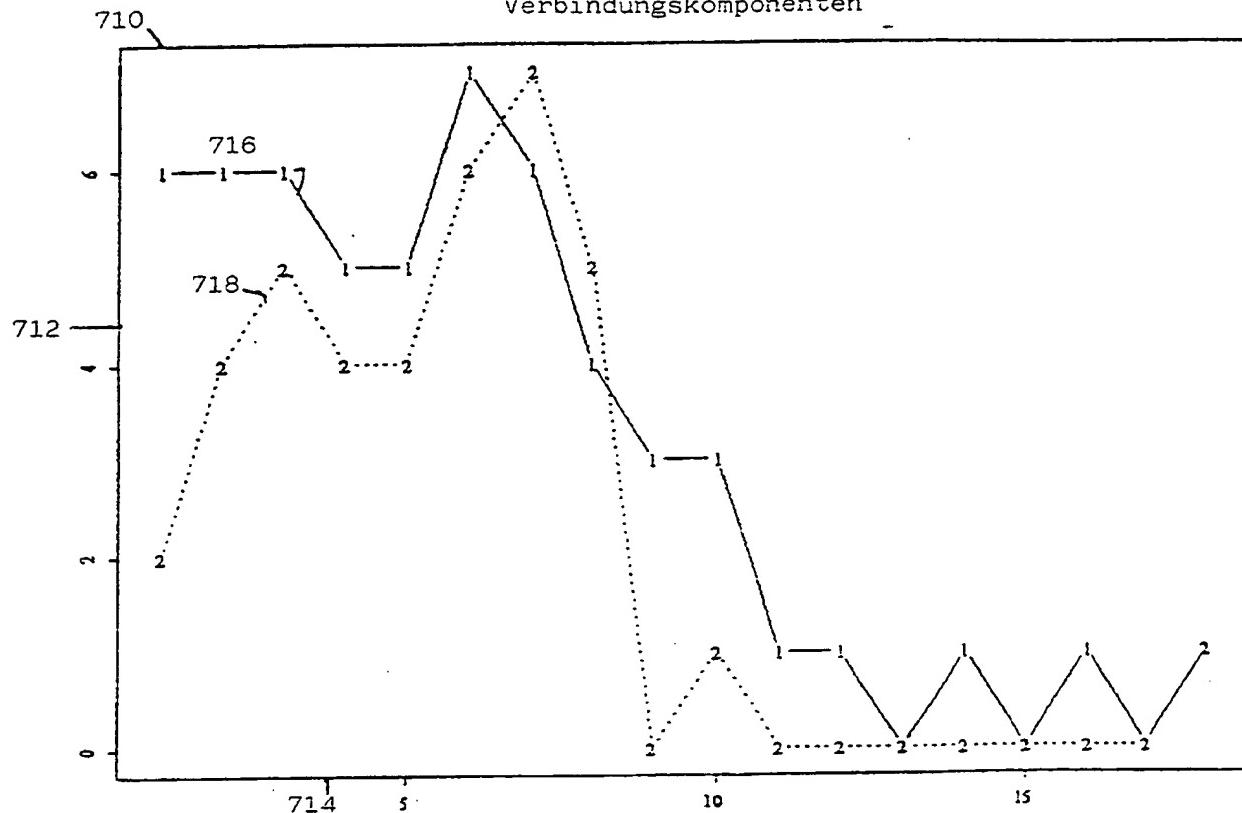
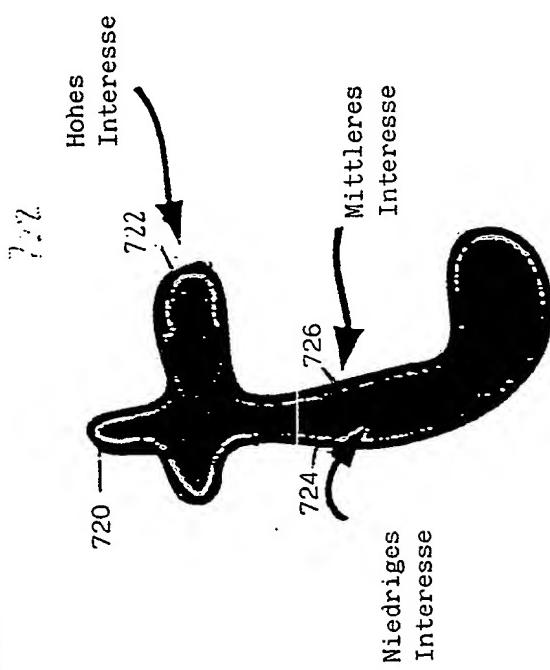


Fig. 73

Interessenoperator

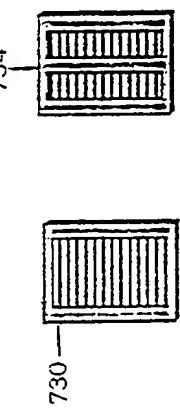
Interessenpunkte



Interessenwerte und Verteilungen ändern
sich für verschiedene Fontgrößen

Fig. 7C

Projektionshistogramm



Zum Unterscheiden

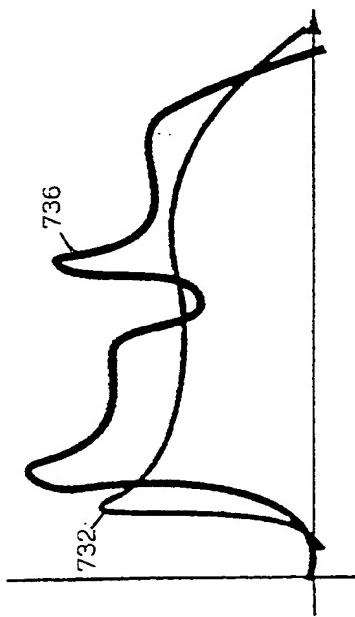


Fig. 7D

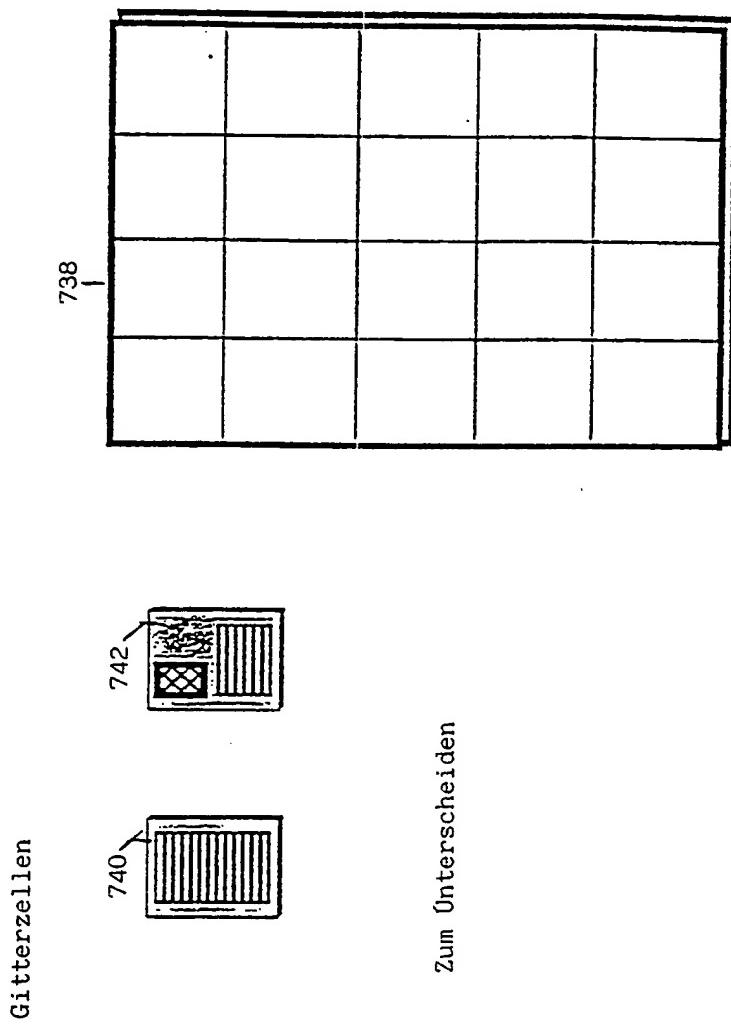


Fig. 7E

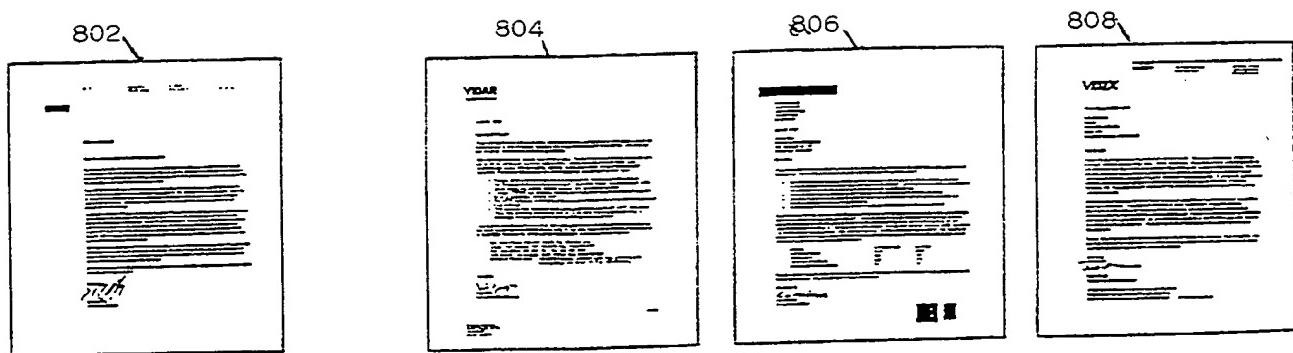


Fig. 8

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLA.

THIS PAGE BLA.